
NLSY79 APPENDIX 5:

SUPPLEMENTAL FERTILITY AND

RELATIONSHIP VARIABLES (2000)

FILE CONTENTS

The supplemental fertility data file, previously referred to as FERTILE, has been renamed “Fertility and Relationship History.” This area of interest contains a variety of constructed and edited variables based on the fertility and marriage histories of respondents, as well as the household record, from the 1979-2000 National Longitudinal Surveys of Youth 1979 cohort (NLSY79). These variables enable users to more easily access the wealth of demographic information provided by the surveys and improve the internal consistency of such data across survey years. The file contains dates of birth, sex, and usual living arrangements for all respondents’ children based on a review of the longitudinal data record. Beginning in 1994, the two-digit IDs of the biological children of the female respondents were added to FERTILITY AND RELATIONSHIP HISTORY, as were separate edit flags for male and female respondents. Also included are created variables that summarize dates of marriage(s), number and type of marriage and/or cohabiting relationships, number of live births and other pregnancy outcomes, spacing between births, spacing between first birth and first marriage, and age of the respondent at the time of the first marriage and key fertility events. The variables included in this file are based upon the youth fertility data as revised in a data cleanup program undertaken in 1982-1983 with additional editing provided at selected subsequent survey points.

1982 DATA QUALITY CHECK AND 1986–90 REVISIONS

Many of the inputs into these fertility variables (specifically month and year of child’s birth) were revised in 1982 in order to maximize internal consistency across years. All of the fertility-related variables in the current file are based on these “revised” data items, with the exception of those variables created directly from the respondent’s household record. The variable R08988.01 (Consistency Of Fertility Data 79-82) specifies for each case whether, after revision, any discrepancy in the fertility history from 1979 to 1982 still remains. In general, when a respondent was interviewed each year from 1979 to 1982, the revised 1982 variables give an accurate picture of the respondent’s fertility history as of 1982. However, some inconsistencies in the data over the period were irresolvable. In such cases, the original 1982 data were left intact and a code was assigned to indicate the nature of the remaining inconsistency. A code of “1” means that the 1982 data are consistent with previous survey years, “2” indicates that a dating error remains, and “3” means that an error in the number of children still remains as of 1982. When an error on dating and on number of children occurred simultaneously, the respondent was coded as having an error on number of children since this type of misreport was considered to be more serious.

Further edits to children’s dates of births for female respondents occurred in 1986 as a result of preparation for the first round of the new child-mother data tape, and in 1987 for male respondents. In 1989, attention was given to correcting subsequent inconsistencies for both males and females. With each successive survey round, an effort is made to fill in previously missing values on children’s birth dates and to include children the respondent has previously failed to report. Since 1990, additional information collected on the children of the female NLSY79 respondents and released on the “Children/Young Adults of the NLSY79” file, has been used to help reconcile inconsistent information for those respondents in the Fertility and Relationship History file. In general, the quality of the fertility record for the female respondents is superior to that reported for the male respondents.

1994 DATA QUALITY CHECK

Beginning in 1994, the supplemental fertility file includes several variables not available for earlier years. For each child of a female respondent, there is now a two-digit identification number variable, which will allow users to more accurately link data from the fertility file with data from the NLSY79 Child file. There is also a comprehensive edit flag for female respondents, allowing users to know which female respondents have had changes made to their fertility record for each survey year compared to previous data releases and what the nature of the change is. Detailed information about the coding categories is provided in the last section of this document. For male respondents, three types of edit flags are provided: two which show the extent of discrepancies between the most recent fertility record available and the current CRF data, and one which indicates cases edited to correct birth order.

As part of the preparation of the 1994 Fertility and Relationship History file, a major data reconciliation was undertaken, comparing the birth records of the female NLSY79 respondents across years. As a result, users may notice discrepancies in these variables across time. It is important for the user to understand that when a date of birth is corrected, we do not change the data for earlier points in time. Thus, there may be inconsistencies in the dates of birth and ages of specific children between the 1994 data and earlier or subsequent reports.

1993 VARIABLE CONSTRUCTION

Although NLSY79 respondents were interviewed in 1993, the Fertility and Relationship History area of interest originally did not contain constructed variables for the 1993 survey year. Data collected in 1993 were used in the 1994 data reconciliation, and some information, such as dates of birth and death, were incorporated into the 1994 or later variables where appropriate. As part of the data work for the NLSY79 2000 data release, however, these 1993 variables were added to the Fertility and Relationship History area of interest.

The set of variables constructed for 1993 is similar to the sets created for other years in which the full battery of pre- and post-natal questions were not asked, such as 1989 or 1991, in that the pregnancy history variables were not created. Dates of birth, gender, and usual residence are constructed for all children, and 2-digit IDs are provided for the children of female respondents. As with the 1994 data, information for the male respondents was not examined as closely as was the information for female respondents. There is a detailed edit flag for the female respondents, as well as the three edit flags for the male respondents. The marriage history variables were also constructed for 1993.

Although the pattern of data evaluation and the creation of edit flags follow that of the 1994 data reconciliation, data comparisons were done only with data from the 1992 survey and earlier. Because of this approach to the data reconciliation and variable construction, the transition from 1993 to 1994 will not be seamless. Users are always advised to use Fertility and Relationship History data from the most recent survey in which a respondent was interviewed.

1979-2000 RELATIONSHIP HISTORY VARIABLES

As an outgrowth of research funded by the National Institute of Child Health and Human Development (NICHD), a series of cross-sectional relationship history variables has been added to the Fertility and Relationship History area of interest. Survey staff carefully examined the names and relationships of household members as reported in the household roster, as well as in the marriage history information collected in various rounds. An attempt was made to identify all

cohabiting partners listed in the household record at any point and to combine this information with the marriage data reported by respondents. In this way, the number of spouse/partners reported across survey years was identified.

Two variables per survey year have been constructed for all interviews through 2000. For each survey, the first variable indicates the total number of spouses/partners a respondent has ever been known to have. The second variable reports the respondent's relationship to the current or most recent spouse/partner. The possible relationship codes are spouse (1), partner (33) or other (36); respondents with no known spouses or partners are coded -999. The code of "other" is used for cases where someone is listed in the household record of a given survey as, for example, an "other non-relative" but is listed as either a partner or a spouse in the preceding or subsequent survey.

These two variables can be used in conjunction to establish a numeric ID for the current or most recent spouse/partner for any given year. The value of the first variable (0-9 as of 2000) is the first digit of the ID, and the value of the second variable (1, 33, or 36) is the remainder. The resulting number indicates the sequential order of the spouse/partner in the respondent's relationship history and the respondent's relationship to that person.

Users should note that the total count of spouses and partners may be understated, because these variables are based on information reported on the interview date. A spouse or partner may have appeared between survey rounds but not have been present at any survey point. Early examination suggests that this applies to only a modest proportion of cases. In some instances, identification of spouses who were present only between rounds may be possible by using the NLSY79 marriage history, as well as the marriage transition information available at each survey point.

PUBLISHED REPORTS

Since 1982, the NLSY79 fertility data have been collected with the support of funding from the National Institute of Child Health and Human Development (NICHD). The 1982 data quality check was also completed under the auspices of NICHD. A comprehensive description of the evaluation procedures used in revising the data, as well as a variety of tabular and multivariate analyses, can be found in the reports entitled "Fertility-Related Data in the 1982 National Longitudinal Survey of Work Experience of Youth: An Evaluation of Data Quality and Some Preliminary Analytical Results" and "Evaluation of Fertility Data and Preliminary Analytical Results from the 1983 (5th round) Survey of the National Longitudinal Surveys of Work Experience of Youth," both prepared by Frank L. Mott, Center for Human Resource Research. The latter report also includes a detailed evaluation of the NLSY79 abortion data. Additional tables referencing the 1986 data can also be found in "Selected Tables: National Longitudinal Surveys of Youth Cohort, May 1987." Evaluations of the marital history data are provided in R. Jean Haurin, "Inconsistencies in the NLSY Marital History Data - 1986 Supplemental Fertility File" and "Marriage and Childbearing of Adults: An Evaluation of the 1992 National Longitudinal Survey of Youth." These reports are available from:

NLS User Services
Center for Human Resource Research
921 Chatham Lane, Suite 100
Columbus, Ohio 43221-2418
(614) 442-7366; Fax (614) 442-7329

Questions regarding the nature of the fertility data should be directed to Canada Keck, who can be reached in writing at the above address, via e-mail at keck@postoffice.chrr.ohio-state.edu, or by phone at (614) 442-7300. Inquiries relating to the purchase of the data should be addressed to the NLS User Services staff at the above address and phone number.

DATA DESCRIPTION

The current Fertility and Relationship History file of 1266 variables includes a small set of fertility and relationship items for all respondents for 1979-1981 and a more extensive set of variables for 1982-2000. These variables include the marriage, relationship and fertility histories of both male and female respondents and the pregnancy histories of the females. The 2000 Fertility and Relationship History file contains what we believe to be the most accurate information for the female respondents as of that survey point. Information about the biological children of the male respondents was not as closely scrutinized as was the females. Users should note several caveats with regard to the creation of specific variables.

- (1) The fertility data of male respondents has typically been less closely scrutinized than that of the female respondent. Beginning in 1993, separate edit flags for males and females have been added to the Fertility and Relationship History record.
- (2) Users may notice discrepancies in dates of birth or gender variables across time. These discrepancies arise as part of the data reconciliation process. Occasionally, a child who is initially reported as a biological child has been later found to be a stillbirth or a non-biological child and removed from the fertility record. It is important for the user to understand that when a date of birth or other information is corrected, we do not change the data for earlier points in time. Thus, there may be inconsistencies in the dates of birth and ages of specific children, or the total number of children, between the current fertility record and earlier reports.
- (3) Variables indicating the number of children or the age of the youngest child in the household refer to the respondent's biological, adopted, or step-children present in the household at the time of the interview. These variables are created by cycling through the household record for the given survey year. The variable titles and labels were historically adjusted for the 1998 data release to make this distinction more apparent.
- (4) Variables relating to the female pregnancy histories such as number of pregnancies, number of miscarriages/stillbirths, month and year began first pregnancy, age began first pregnancy, and outcome of first pregnancy have valid values only for female respondents interviewed at the survey year in question and who were also interviewed at the time of the 1982 and 1983 surveys when full retrospective pregnancy histories were collected. All male respondents as well as female respondents not interviewed in both 1982 and 1983 are coded as a "4" on these variables since a complete pregnancy profile is unavailable. Beginning with 1992, miscarriages and stillbirths are collapsed into a single code ("2") on the variable "Outcome of First Pregnancy."

Confidential abortion reports were collected in 1984, 1986, 1988, 1990, 1992, 1994, 1996, 1998, and 2000. This information has been incorporated into the creation of the pregnancy-related variables. For that small subset of female respondents for whom full

pregnancy histories are unavailable, some will have full abortion data if they were interviewed in 1984. Thus, there are smaller numbers of respondents with a code of “-4” on the variable for number of abortions than on the other pregnancy variables.

Current pregnancies are included in the count of the number of pregnancies as of a given survey date, and twins/triplets represent a single pregnancy. Where questionnaire items for the beginning date of the first pregnancy are unavailable and the outcome was a live birth, 9 months are subtracted from the child’s birth date to obtain the beginning date of the first pregnancy. Where the outcome of the first pregnancy was an abortion reported only in one of the confidential reports, 3 months are subtracted.

- (5) Beginning with the release of the 1985 marriage variables, an effort has been made to reconcile marriage dates with the key variables for current marital status made available in the “Key Variables” area of interest. For approximately 100 cases there are inconsistencies in the marriage histories over time, with some respondents changing their marital status from ever-married to never-married or vice-versa based on the marital status change item provided on the information sheet. Where the change was made very early in the longitudinal record and the respondent continued to verify the changed status in subsequent surveys, the marriage variables for 1985 and subsequent years have been altered accordingly and will differ from the marriage variables provided earlier. Where a change is recorded from never-married to ever-married using the information sheet item only, marriage dates were not collected and thus the respondent is missing information (or coded “-3”) on the date of marriage. For all survey years, a marriage is considered to have ended only if the respondent reports a change to widowhood or divorce.
- (6) All age variables referenced to events are constructed with the original date of birth of the respondent provided at the 1979 survey (R00003., R00004. and R00005). These variables were used to define a respondent’s eligibility to be included in the NLSY79 sample.
- (7) The variables indicating months between first marriage and first birth have traditionally ranged from negative to positive numbers, with specialized codes to indicate non-interview, no first child, no first marriage, and missing date information. Beginning with the 1998 data release, these data indicating months between first marriage and first birth for each FERTILE record 1982–2000 have been converted to all positive numbers. These variables are paired with a flag variable that indicates whether the first marriage occurred before or after the first birth. Both variables have been assigned new reference numbers. The specialized codes for non-interviews have been dropped; however, the other specialized codes have been retained. The original versions of these variables **do not** appear in the public release.
- (8) The newly added relationship history variables help users track the number of spouse/partners identified through our data collection process. However, not all cohabiting partners may have been reported as such, and cohabiting partners could also have entered and exited households between survey rounds and thus be unavailable for identification.

CODEBOOK CATEGORIES

The following variables have special coding specifications that users need to be aware of when using the 1994–2000 supplemental fertility data.

R50882. Edit Flag for Female Respondents 1994 Survey
 R51735. Edit Flag for Female Respondents 1996 Survey
 R64871. Edit Flag for Female Respondents 1998 Survey
 R70149. Edit Flag for Female Respondents 2000 Survey

This edit flag is a general code that indicates the status of the MOTHER'S fertility record, and the indicated changes do not specify which child was affected.

- 0 = Consistent with previous supplemental fertility file records
- 1 = A child made younger
- 2 = A child made older
- 3 = Corrected previously missing information
- 4 = Information inconsistent with previous supplemental fertility; previous supplemental fertility information used
- 5 = Information inconsistent with previous supplemental fertility; mother's current information accepted reluctantly
- 6 = IDs assigned out of birth order
- 7 = Discrepancy between CRF and FERTILE; current FERTILE will be consistent with previous supplemental fertility, but CRF preserved for next survey round
- 8 = Child removed from FERTILE; incorrectly recorded non-biological child
- 9 = Child removed from FERTILE; incorrectly recorded pregnancy loss
- 10 = IDs assigned out of birth order AND made a child younger
- 11 = IDs assigned out of birth order AND made a child older
- 12 = Gender changed from previous supplemental fertility
- 13 = Data from mother inconsistent; one child deleted, another added
- 14 = IDs out of birth order AND discrepancy between FERTILE and CRF (CRF preserved for next survey round)
- 15 = Child removed from FERTILE; incorrectly recorded pregnancy loss AND changed gender on another child from previous FERTILE
- 16 = Made current supplemental fertility consistent with current Child Supplement (new child)
- 17 = Sex missing from current CRF (new child); used information from current CS
- 18 = Child assessed in current round but inexplicably missing from current CRF; added to current supplemental fertility with residence information from HHR if possible
- 19 = Hand edited date of death
- 20 = Day of birth ONLY discrepancy between previous supplemental fertility and current CRF; unedited CRF day used
- 21 = Child assessed in current survey round but mother is a noninterview
- 22 = Mother added surprise older child; IDs out of birth order
- 23 = Incorrect code of 99 (deleted) generated for deceased/adopted out child; information corrected
- 24 = Duplicate date of birth of existing child on CRF; edited to reflect previous supplemental fertility
- 25 = Non-biological child not previously on FERTILE deleted from current CRF
- 26 = Incorrect code of 99 (deleted) generated for live child; status corrected and residence taken from HHR if possible
- 27 = Incorrect HH flag generated by CAPI for deceased/adopted out child; information corrected (New code in 1996)
- 28 = Residence missing from CRF; information from fertility section and/or HHR used (New code in 1996)

- 29 = Partial interview; used previous supplemental fertility file; residence coded from HH record if possible (New code in 1998)
- 30 = Incorrect code of 8 (deceased) generated for live child; status corrected and residence taken from HHR if possible (new code in 1998)
- 31 = Information corrected based on YA respondent correcting birth, age not affected
- 32 = Information corrected based on YA respondent correcting birth, age made older
- 33 = Information corrected based on YA respondent correcting birth, age made younger
- 34 = Child removed from FERTILE, added through interviewer error but not caught previously
- 35 = Residence coded based on information from HHR

There are three edit flags for MALE respondents for 1994–2000. The coding scheme for these flags appears in the codebook.